

Gaussian Boson Sampling

Craig S. Hamilton,^{1,*} Regina Kruse,² Linda Sansoni,² Sonja Barkhofen,² Christine Silberhorn,² and Igor Jex¹

¹*FNSPE, Czech Technical University in Prague, Brřhová 7, 119 15, Praha 1, Czech Republic*

²*Integrated Quantum Optics, Universität Paderborn, Warburger Strasse 100, 33098 Paderborn, Germany*

Boson sampling has emerged as a tool to demonstrate the difference between quantum and classical computers and has piqued the interest of experimentalists and theoreticians. Its main advantage is that it does not require full universal control over the quantum system, which favours current photonic experimental platforms for implementation. However, these platforms suffer from a scaling problem because of probabilistic state generation. We solve this problem by devising a boson sampling protocol using Gaussian states and show that this is a more general type of photonic boson sampler. First we derive a new result that relates the probability of a photon pattern generated by a general Gaussian state and show that it depends upon a matrix function called the hafnian. This leads to an experimentally feasible photonic boson sampler exponentially more likely to generate high photon numbers, going far beyond the limitations of previous scenarios.

Boson sampling has sparked the imagination of theorists and experimentalists since it was introduced by Aaronson and Arkhipov (AABS) [1]. It demonstrates the power of quantum computation over classical computation and provides evidence against the Extended Church-Turing (ECT) theorem, without the need for the full power of universal quantum computation. In photonic boson sampling, single photon Fock states are launched into a multimode interferometer where, due to bosonic statistics, the probability of the output distribution of photons depends upon the permanent of a matrix characteristic for the interferometer. The permanent is in the #P complexity class [2] therefore this distribution is difficult to sample from unless certain computational complexity classes are equivalent, which would have serious consequences for complexity theory.

After this theoretical advance various experimental groups performed the first demonstrations of the protocol [3–6]. To simplify the task of generating many single photons with probabilistic sources a new scheme of boson sampling, known as scattershot boson sampling (SBS) was formulated [7] and implemented [8, 9]. Here, photons from Two Mode Squeezed States (TMSS) enter a linear interferometer and a detection setup. The detected photons ‘herald’ the photons that enter the interferometer, which undergo a linear transformation and are then measured at the output. However, postselected Fock states are generated probabilistically and therefore the scaling to high photon numbers has a intrinsic exponential cost in these schemes.

It is therefore valuable from an experimental perspective to investigate other quantum photonic states that can produce problems in the same complexity class as Fock states with simple dynamics. One such family of states are Gaussian states [10]; easy to generate experimentally and have been studied in great detail [11]. Previous work used postselected single photons from Gaussian states [7] that then enter the interferometer and are

sampled. However, from the point of view of Gaussian states this is restricted to a very specific interferometric setup or, alternatively, the sampling is done from Fock states. Also, in [12] the authors show that sampling from a multimode thermal state is in BPP^{NP} (not as hard as AABS). This leaves open the question of a sampling from a general Gaussian state.

In this letter we answer this question of sampling photons from a general, non-classical Gaussian state with a new protocol we call Gaussian Boson Sampling (GBS). Our definition of GBS is more general than what has been previously studied, as we consider Single Mode Squeezed States (SMSS) entering an arbitrary linear interferometer and sample the output with single photon detection. We first derive a new result that shows the probability of a photon output distribution from a general Gaussian state can be written in terms of a matrix function, the *hafnian*, which is a more general matrix function than the permanent and this is the first use of it in a quantum optics setting. We then show that GBS with SMSS is *exponentially* more likely to produce high number of photons when compared to SBS, a substantial experimental advantage and one that puts photonic boson sampling within the grasp of today’s technology.

This new regime of photonic boson sampling is currently the most general form available, as all others can be derived from our model (starting with Gaussian states and incorporating postselection into the dynamics). This can be most readily seen from SBS but is also present in other mentioned forms.

Output photo-counts from a Gaussian state

Photonic boson sampling involves sending single photon Fock states into a linear interferometer, described by a matrix T , which transforms the M modes. The probability of measuring a certain pattern of photons $\hat{n} = \bigotimes_j^M n_j |n_j\rangle \langle n_j|$ (n_j photons in output mode j) from M modes of a quantum state is $\text{Pr}(\bar{n}) = \text{Tr}[\hat{\rho} \hat{n}]$. For boson sampling from Fock states $\text{Pr}(\bar{n})$ depends upon the

*Electronic address: hamilcra@fjfi.cvut.cz

permanent of a matrix [13],

$$\Pr(\bar{n}) = \frac{|\text{Perm}(T_S)|^2}{\bar{n}! \bar{m}!}, \quad (1)$$

where \bar{m} is the input photon pattern, $\bar{n}! = n_1! n_2! \dots n_M!$ and T_S is a submatrix of the linear transformation that depends upon where the photons enter and exit the interferometer.

Here we derive a new expression for $\Pr(\bar{n})$ from a Gaussian input state. These states are characterised solely by a $2M \times 2M$ covariance matrix σ and a displacement vector d [14],

$$\sigma_{ij} = \frac{1}{2} \langle \{\hat{\xi}_i, \hat{\xi}_j\} \rangle, \quad d_j = \langle \hat{a}_j \rangle,$$

where $\hat{\xi}$ run over all $\hat{a}_j, \hat{a}_j^\dagger$ (in this paper we will assume $d_j = 0 \forall j$). The details of this derivation are given in Appendix A.

Using phase space methods (similar to [15, 16] and [12]), $\Pr(\bar{n})$ becomes the integral of the Q- and P-functions of the state and operator,

$$\Pr(\bar{n}) = \pi^M \prod_{j=1}^M \int d^2 \alpha_j Q_{\hat{\rho}}(\alpha) P_{\bar{n}}(\alpha). \quad (2)$$

where $d^2 \alpha_j = d\alpha_j d\alpha_j^*$, $Q_{\hat{\rho}}$ is the Q-function of the state and $P_{\bar{n}}$ is the P-function corresponding to the operator $\hat{n} = \bigotimes_{j=1}^M \hat{n}_j$. This analysis leads to,

$$\Pr(\bar{n}) = \frac{1}{\bar{n}! \sqrt{|\sigma_Q|}} \prod_{j=1}^M \left(\frac{\partial^2}{\partial \alpha_j \partial \alpha_j^*} \right)^{n_j} e^{\frac{1}{2} \alpha_v^\dagger A \alpha_v} \Bigg|_{\alpha_v=0}, \quad (3)$$

where $\sigma_Q = \sigma + \mathbb{I}_{2M}/2$, $\alpha_v = [\alpha_1, \dots, \alpha_M, \alpha_1^*, \dots, \alpha_M^*]$ and

$$A = \begin{pmatrix} 0 & \mathbb{I}_M \\ \mathbb{I}_M & 0 \end{pmatrix} [\mathbb{I}_{2M} - \sigma_Q^{-1}]. \quad (4)$$

A can be divided into four block matrices, shown in figure 1, which is a result of the initial structure of σ .

For simplicity we focus on $n_j = \{0, 1\}$ (we will deal with $n_j \geq 2$ in Appendix B) for a total of N photons and $2N$ derivatives (for $\partial \alpha_j, \partial \alpha_j^*$). The N indices of the photons' mode-position are written in a permutation vector μ of length $2N$ with entries j and $j + M$ per photon. The $2N$ derivatives select the rows/columns of A where the photons were measured; the other rows/columns can be discarded. This is illustrated in figure 1, where the intersection of the rows and columns where a photon was detected, highlighted in blue, form the entries of the submatrix, A_S .

The expansion of the $2N$ derivatives leads to a summation over all perfect matching permutations (PMP)

$$A^{2M \times 2M} = \begin{pmatrix} B & C \\ C^t & B^* \end{pmatrix} \longrightarrow A_S^{2N \times 2N}$$

FIG. 1: Construction of submatrix A_S from A . Also shown is the structure of A , which can be divided into 4 block matrices.

[17, 18] of the vector μ . For a general matrix A_S this is,

$$\Pr(\bar{n}) = \frac{1}{\bar{n}! \sqrt{|\sigma_Q|}} \sum_{\mu \in \text{PMP}} \prod_{j=1}^M A_{S_{\mu(2j-1), \mu(2j)}}. \quad (5)$$

The sum over all PMP is exactly the hafnian of A_S , as defined by Caianiello [19, 20]. Finally we arrive at,

$$\Pr(\bar{n}) = \frac{1}{\bar{n}!} \frac{\text{Haf}(A_S)}{\sqrt{|\sigma_Q|}}. \quad (6)$$

This new result relates the probability of a photon pattern \bar{n} from a general Gaussian state to the hafnian of a matrix that characterises that state. This deals with any covariance matrix of a Gaussian state, with both squeezed and thermal contributions and can also include loss and noise dynamics, which are described by Gaussian operations.

We use this result to develop a boson sampling protocol for Gaussian states, with squeezing contribution only (this situation corresponds to the matrix A in figure 1 having submatrix $C = 0$).

Gaussian boson sampling with squeezed states

As the hafnian is in the #P-complete complexity class [2], we can use (6) to devise a quantum sampling problem akin to AABS. Whereas the permanent counts the (weighted) number of perfect matchings in a bipartite graph, the hafnian counts the number of perfect matchings in a general graph (not necessarily bipartite). Thus the hafnian is a more general function than the permanent, which is encapsulated in the formula [21],

$$\text{Perm}(G) = \text{Haf} \begin{pmatrix} 0 & G \\ G^t & 0 \end{pmatrix}, \quad (7)$$

where G is the adjacency matrix of the graph (see Appendix C for decompositions of the hafnian). This means that any algorithm or black box that can accurately calculate the hafnian could also calculate the permanent, which is known to be #P-hard even to approximate [1]. Currently, there is no algorithm to efficiently approximate the hafnian [22, 23].

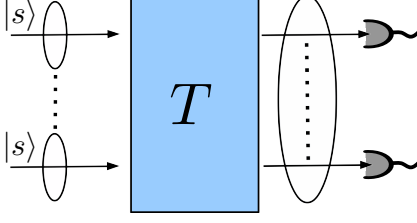


FIG. 2: Gaussian Boson Sampling. SMSS enter a linear interferometer, T , and at the output we measure the multimode squeezed state in the Fock state basis. The probability of a given pattern \bar{n} is given by (11).

We illustrate GBS with the scenario shown in figure 2. SMSS enter an M -mode linear interferometer, described by a Haar random unitary T , with all modes being measured at the output. The squeezing transformation is described by,

$$S = \begin{pmatrix} \bigoplus_{j=1}^M \cosh r_j & \bigoplus_{j=1}^M \sinh r_j \\ \bigoplus_{j=1}^M \sinh r_j & \bigoplus_{j=1}^M \cosh r_j \end{pmatrix}, \quad (8)$$

which is divided into 4 block diagonal matrices [24] and r_j is the squeezing parameter of the j^{th} mode. The covariance matrix at the output of the interferometer is [25],

$$\sigma = \frac{1}{2} \begin{pmatrix} T & 0 \\ 0 & T^* \end{pmatrix} S S^\dagger \begin{pmatrix} T^\dagger & 0 \\ 0 & T^t \end{pmatrix}, \quad (9)$$

and we arrive at $A = B \oplus B^*$ where,

$$B = T \left(\bigoplus_{j=1}^M \tanh r_j \right) T^t, \quad (10)$$

The probability of measuring \bar{n} (0 or 1 photon only per mode) is then,

$$\Pr(\bar{n}) = \frac{1}{\sqrt{|\sigma_Q|}} |\text{Haf}(B_S)|^2, \quad (11)$$

where the matrix B_S is the submatrix that only has the rows/columns remaining where a photon was detected (see figure 1). Below we discuss other requirements of the GBS protocol.

Random Haar Unitary & Dilute Boson Sampling

A condition for AABS is that the transformation T is drawn from the Haar random unitary family, which ensures that the submatrix, T_S , has elements that are randomly distributed according to a complex normal distribution [1].

In GBS it is not T that is sampled but B (10). If we assume $r_j = r, \forall j$ then $B = \tanh r \times T T^t$, and $T T^t$ is from the Circular Orthogonal Ensemble (COE) of random matrices (if T is a Haar random unitary). A similar result [26] shows that a submatrix of B will have entries that are i.i.d. from the complex normal distribution if the $m \times m$ submatrix is small enough $m = O(\sqrt{M})$.

If we pump $K (< M)$ modes this means that B is a rank K matrix. It is known that matrix rank determines the complexity of calculating the permanent [27]. Thus, we will assume a similar result for hafnians which means that to detect N photons we will require to sample from (at least) a rank N matrix and therefore require that $K \geq N$ SMSS at the input.

AABS requires ‘dilute’ positions of photons at the output i.e. a sufficiently low probability of ≥ 2 photons in the same mode, as this will maximise the rank of the matrix T_S . To demonstrate that GBS is capable of dilute BS we look at the output covariance matrix of an individual mode. We start with N SMSS, of the total $M = N^2$ modes, with an average number of photons per SMSS equal to unity. The Haar-random transformation rotates the average photon number into all modes equally (on average), making the mean number of photons per mode $\approx 1/N$. Taking a single-mode subsystem (tracing over all other modes) yields a thermal state with average photon number $\approx 1/N$ and the ratio of photon number probability is $\Pr(2)/\Pr(1) \approx 0.1$, sufficient for dilute BS.

What are we sampling in GBS?

In AABS there is a fixed number photons, N , that enter and exit the interferometer. Thus the set of probabilities (of size $C_N = \binom{M+N-1}{N}$) sampled are,

$$\{p_1, p_2, \dots, p_{C_N}\}_N = \{P_N\}, \quad (12)$$

where $p_j = \Pr(\bar{n})$, $\sum_j p_j = 1$ is the probability of a certain output pattern, given by (1), and $\{P_N\}$ is the set of all output patterns with N photons. GBS does not have a fixed number of output photons due to the nature of Gaussian states. That means the sampled set is all patterns with N photons ($N \in [0, \infty)$),

$$\begin{aligned} \{ \{p = |\sigma_Q| \}_0, \{p_1, p_2, \dots, p_{C_1}\}_1, \dots, \{p_1, p_2, \dots, p_{C_N}\}_N, \dots \} \\ = \{ \{P_0\}, \{P_1\}, \{P_2\}, \dots, \{P_N\}, \dots \} \end{aligned}$$

and again the probabilities sum to unity. However, dilute boson sampling and squeezing matrix rank (8) mean that we will have to discard certain photon patterns as the complexity of such matrices are in question. We will show how to control the number of photons in our system below.

Number of photon pair events from multiple squeezers

The probability of N Photon Pair Events (PPE) events ($2N$ photons) from a SMSS is given by the negative binomial distribution [28] and using the properties of this distribution the probability of creating N PPE events from $K \times$ SMSS is,

$$P_K(N) = \binom{K/2 + N - 1}{N} \text{sech}^K(r) \tanh^{2N}(r) \\ = \frac{\Gamma(K/2 + N)}{\Gamma(K/2) N!} \text{sech}^K(r) \tanh^{2N}(r). \quad (13)$$

In figure 3 we plot this distribution for 14 SMSS with $r = 0.8814$ (average number of photons per squeezer is 1.) For this distribution the mean number of photons is $K \sinh^2(r)$ and the modal number is $(K - 1) \sinh^2(r)$. In figure 3 we also highlight the probabilities (in red) of 5–7 PPE events (10–14 photons). These events would yield computationally complex outcomes for a large ($14^2 = 196$ modes) interferometer, ‘saturating’ the complexity of the device by reducing the incidence of multiple photons in the same mode and randomising the entries of the sampled matrix. The outcomes deemed ‘too many’ photons may be useable, but this would require extra complexity theory arguments and, experimentally, photon-number discriminating detectors.

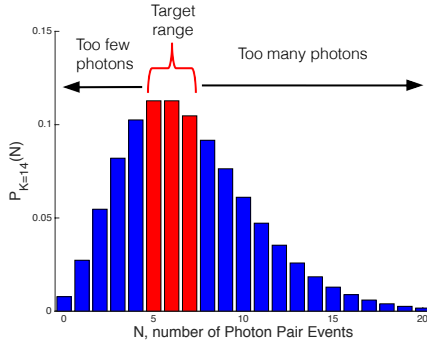


FIG. 3: The negative binomial distribution with $14 \times$ SMSS ($K = 14$) with squeezing parameter $r = 0.8814$.

The red bars are the ‘useful’ photons occurrences, 10–14 photons.

In summary, to perform GBS with N photons an experiment will require an interferometer of size $M = N^2$ and will pump K modes $K \geq N$ modes with squeezed light. We can either operate in a regime where we focus on a single output photon number N , and make that the modal number of the distribution, or where we have a range of output numbers, $[N - c, N]$ (where c is a small integer), that has a high probability to be generated. These choices of photon number ‘saturate’ the complexity of the interferometer. We then measure all modes at the output and analysis results where the photons are found one per mode.

Relationship between SBS and GBS

Figure 4 illustrates the relationship between SBS and GBS. We begin with $2M$ weakly pumped (with equal squeezing parameter r) SMSS connected in parallel that, taken in pairs, create two-mode squeezed states (TMSS) when they enter a phase shifter and beamsplitter described respectively by the transformations,

$$U_{PS} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, U_{BS} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

After creating the TMSS, one mode enters the herald-detection setup while the other mode enters the interferometric setup, T_{SBS} , that linearly transforms the modes. The overall interferometer for the system is $(I \oplus T_{SBS}) \times (\bigoplus_j U_{BS} U_{PS})$ which is a special case of the GBS interferometer shown in figure 2.

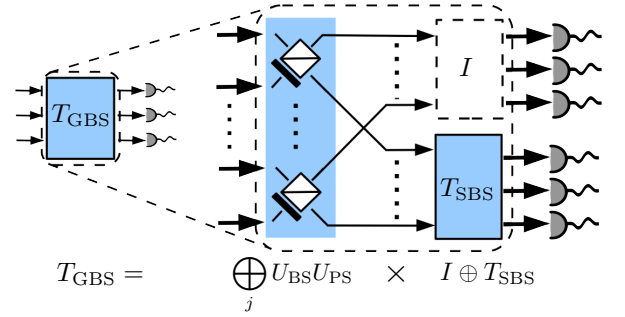


FIG. 4: Relationship between SBS and GBS, showing how SBS is a subset of GBS. First TMSS are created by the pairwise interactions of $U_{BS} U_{PS}$ which then enter the interferometer $I \oplus T_{SBS}$. One half of the modes are used to herald the other half that are transformed by T_{SBS} , which are then detected for SBS.

Randomly, herald detectors fire, generating herald pattern \bar{h} and tell us the input pattern for the modes entering the interferometer T_{SBS} . A necessary requirement for SBS is that we only detect a single photon-pair from each TMSS. The probability to measure \bar{n} at the output depends upon where the photons entered, i.e. \bar{h} . Therefore, the actual probability for SBS is a conditional one,

$$\text{Pr}_{\text{SBS}}(\bar{n}) = \text{Pr}(\bar{n}|\bar{h}) = \frac{\text{Pr}(\bar{n} \cap \bar{h})}{\text{Pr}(\bar{h})}, \quad (14)$$

where we have used Bayes’ theorem to relate the conditional probability of the SBS scheme with a joint probability of the photon pattern $\bar{n} \cap \bar{h}$. $\text{Pr}(\bar{n} \cap \bar{h})$ is nothing more than the probability of measuring the photon pattern $\bar{n} \cap \bar{h}$ from the $2M \times$ SMSS entering the interferometer $(I \oplus T) \times (\bigoplus_j U_{BS} U_{PS})$, in other words GBS. The denominator, $\text{Pr}(\bar{h})$, is the probability of the herald photons to be detected which is also (in this simple system) the probability to generate the N PPE events from M TMSS. This simple statement shows the relationship

between SBS and GBS without appealing to any matrix function. In Appendix D we explore this in more detail by using the results derived in this paper.

SBS uses $K \times \text{TMSS}$ and only retains the occurrences of a single PPE from each squeezer. Thus the probability of N PPE's follows a binomial distribution,

$$P_{\text{SBS}}(N) = \binom{K}{N} \text{sech}^{2K}(r) \tanh^{2N}(r), \quad (15)$$

which always has lower probabilities when compared to GBS, (13). To illustrate this, we compare (13) and (15) for $2K \times \text{SMSS}$ and $r_{\text{SBS}} = r_{\text{GBS}}$. The ratio of these is,

$$\begin{aligned} P_{\text{SBS}}(N)/P_{\text{GBS}}(N) &= \binom{K}{N} / \binom{K+N-1}{N} \\ &= \frac{K!(K-1)!}{(K-N)!(K+N-1)!} \\ &\approx \lim_{N \rightarrow \infty, K > N} \left(\frac{K-N}{K-1} \right)^N. \end{aligned} \quad (16)$$

This ratio scales exponentially in favour of GBS, thereby demonstrating the substantial advantage of our protocol. This is due to the fact that GBS makes use of higher-order PPE's from the *same* individual squeezer (SBS only uses the first event per squeezer), making GBS much more efficient at generating large numbers of photons.

There are two other significant factors in photon generation. First, all generated photons will be used for sampling as none are used for heralding (a factor of 2). Next, the squeezing parameter, which represents the strength of the squeezing process, can be significantly higher in GBS. For SBS $r_{\text{SBS}} \ll 1$, as we only want a single PPE, but for GBS, $r_{\text{GBS}} \approx 1$ as we can use multiple photons from the same squeezer. This is an additional exponential factor in generating photons favouring GBS.

Conclusions

Boson sampling has emerged as a simple yet powerful tool to test quantum physics in the field of computation. Currently, the best protocol for achieving this, SBS, scales badly with the number of photons as it relies upon non-deterministic state generation.

To solve this problem and open up new avenues to realise feasible photonic boson sampling, we have introduced Gaussian boson sampling. We have derived an expression for the output probabilities from a general Gaussian state, showing that it is a computationally difficult problem by relating it to a matrix function called the hafnian. The hafnian is in the complexity class $\#P$ -complete as it is a more general function than the permanent. Physically, this can be seen in the boson sampling framework as SBS being a subset of all possible GBS realisations. Additionally, GBS has a significant advantage to generate high photon numbers, making use of multiple

PPE's from the same squeezed sources, when compared to SBS. We also note that our result, (6), with only thermal contributions, corresponding to $B = 0, C \neq 0$ in figure 1, replicates the results of [12].

These results answer questions in previous work as to the complexity of boson sampling with Gaussian states [1, 12, 29]. Within experimental quantum optics, starting with a Gaussian state, i.e. a squeezed state, and then using linear optics and postselected measurement outcomes is a very common method to create different families of photonic states. Therefore our work naturally includes other boson sampling problems such as those involving Schrödinger cat states and photon added/subtracted states [30–32]. This means that GBS is a more general problem than what has previously been studied. As the types of states used for GBS, multimode squeezed states, also known as cluster states, are universal for quantum computation [33], it seems that GBS is the most general case of photonic boson sampling. Also, due to the symmetry of quantum mechanics we can invert the problem we look at here, Gaussian states input to an interferometer with Fock-basis measurements, and use the same result to explain Fock states input to an interferometer with Gaussian-basis measurements (see open problem 4 in [1]). Our result, (6), can be used to describe other optical experiments and therefore may find other uses outside of boson sampling experiments.

This formalism allows us to handle the main source of noise in photonic systems, photon loss and dark-counts, in a very natural way as both are Gaussian operations. This makes lossy GBS easy to deal with compared to lossy SBS [34], though it remains an open question how much loss we are able to tolerate and retain the $\#P$ complexity of the scheme (as opposed to the BPP^{NP} complexity of thermal states.) Also to be noted is that we conjecture that approximating the hafnian of matrices of random i.i.d. complex variables is $\#P$ -hard, using the same reasoning that calculating the hafnian is difficult (it is more general than the permanent).

Most importantly, we have shown by replacing the single photon input with squeezed light a new class of photonic boson samplers can be reached. GBS demonstrates a way to achieve a photonic boson sampler with a significant number of photons to show the supremacy of quantum computers.

Acknowledgements

C.S.H. and I. J. received support from the Grant Agency of the Czech Republic under grants No. GACR 13-33906S, RVO 68407700. This work has received funding from the European Union's Horizon 2020 research and innovation program under the QCHIP project Grant No. 641039.

The authors would like to thank A. Björklund and S. Rahimi-Keshari for useful comments.

-
- [1] S. Aaronson and A. Arkhipov, *Theory of Comput.* **9**, 143 (2013).
- [2] L. Valiant, *Theoretical computer science* **8**, 189 (1979).
- [3] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, *Science* **339**, 794 (2013).
- [4] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Nature Photon* **7**, 540 (2013).
- [5] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, et al., *Science* **339**, 798 (2013).
- [6] N. Spagnolo, C. Vitelli, M. Bentivegna, D. J. Brod, A. Crespi, F. Flamini, S. Giacomini, G. Milani, R. Ramponi, P. Mataloni, et al., *Nature Photonics* **8**, 615 (2014).
- [7] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, *Physical Review Letters* **113**, 100502 (2014).
- [8] J. Carolan, J. D. A. Meinecke, P. J. Shadbolt, N. J. Russell, N. Ismail, K. Wörhoff, T. Rudolph, M. G. Thompson, J. L. O'Brien, J. C. F. Matthews, et al., *Nature Photonics* **8**, 621 (2014).
- [9] M. Bentivegna, N. Spagnolo, C. Vitelli, F. Flamini, N. Viggianiello, L. Latmiral, P. Mataloni, D. J. Brod, E. F. Galvão, A. Crespi, et al., **1** (2015).
- [10] B. L. Schumaker, *Physics Reports* **135**, 317 (1986).
- [11] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, *Rev. Mod. Phys.* **84**, 621 (2012).
- [12] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, *Physical Review Letters* **114**, 060501 (2015).
- [13] S. Scheel, arXiv **quant-ph** (2004), quant-ph/0406127v1.
- [14] A. Ferraro, S. Olivares, and M. G. A. Paris, arXiv **quant-ph** (2005), quant-ph/0503237v1.
- [15] V. V. Dodonov, O. V. Man'ko, and V. I. Man'ko, *Physical Review A. Atomic* **49**, 2993 (1994).
- [16] V. V. Dodonov, O. V. Man'ko, and V. I. Man'ko, *Physical Review A* **50**, 813 (1994).
- [17] D. Callan, arXiv **math.CO** (2009), 0906.1317v1, URL <http://arxiv.org/abs/0906.1317v1>.
- [18] PMP are permutations μ that obey 1) $\mu(2j-1) < \mu(2j)$ & 2) $\mu(2j) < \mu(2(j+1)) \forall j$.
- [19] E. R. Caianiello, *Il Nuovo Cimento* **10**, 1 (1953).
- [20] E. R. Caianiello, *Combinatorics and Renormalization in Quantum Field Theory* (W. A. Benjamin, Inc., 1973).
- [21] H. Minc, *Permanents* (Addison-Wesley, 1978).
- [22] P. Sankowski, in *Annual Symposium on Theoretical Aspects of Computer Science* (Springer, 2003), pp. 427–438.
- [23] A. Björklund, in *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms* (SIAM, 2012), pp. 914–921.
- [24] $\bigoplus_{j=1}^N A_j = A_1 \oplus A_2 \dots \oplus A_N = \text{diag}(A_1, A_2, \dots, A_N)$.
- [25] R. Simon, N. Mukunda, and B. Dutta, *Physical Review A (Atomic)* **49**, 1567 (1994).
- [26] T. Jiang, *Journal of Mathematical Physics* **50**, 063302 (2009).
- [27] A. I. Barvinok, *Mathematics of Operations Research* **21**, 65 (1996).
- [28] J. M. Hilbe, *Negative Binomial Regression* (Cambridge University Press, 2011).
- [29] J. Huh, G. G. Guerreschi, B. Peropadre, J. R. McClean, and A. Aspuru-Guzik, *Nature Photonics* **9**, 615 (2015).
- [30] P. P. Rohde, K. R. Motes, P. A. Knott, J. Fitzsimons, W. J. Munro, and J. P. Dowling, *Physical Review A* **91**, 012342 (2015).
- [31] J. P. Olson, K. P. Seshadreesan, K. R. Motes, P. P. Rohde, and J. P. Dowling, *Physical Review A* **91**, 022317 (2015).
- [32] K. P. Seshadreesan, J. P. Olson, K. R. Motes, P. P. Rohde, and J. P. Dowling, *Physical Review A* **91**, 022334 (2015).
- [33] N. C. Menicucci, P. V. Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, *Phys. Rev. Lett.* **97**, 110501 (2006).
- [34] S. Aaronson and D. J. Brod, *Physical Review A* **93**, 012335 (2016).
- [35] S. M. Barnett and P. Radmore, *Methods in Theoretical Quantum Optics* (OUP, 1996).
- [36] L. Comtet, *Advanced Combinatorics* (D. Reidel Publishing Company, 1974).
- [37] M. Hardy, *Electron. J. Combin* **13**, 13 (2006).
- [38] M. L. Mehta, *Random Matrices* (Elsevier, 2004).
- [39] R. P. Stanley, *Enumerative Combinatorics Vol. 1* (Cambridge University Press, 1997).

Appendix A: Derivation of Equation (6)

Here we explain in more detail the derivation of (6) in the main text. The Q-function of a Gaussian state is the convolution of the Wigner function with another Gaussian [35],

$$Q_{\hat{\rho}}(\alpha) = \frac{1}{\sqrt{|\pi\sigma_Q|}} \exp\left[-\frac{1}{2}\alpha_v^\dagger \sigma_Q^{-1} \alpha_v\right] \quad (\text{A1})$$

where σ_Q and α_v are defined in the main text. Note that the σ_Q here corresponds to the system modes that are measured. If we do not measure a mode then we trace over those system modes by removing the rows and columns (two of each) from σ_Q corresponding to that mode. The P function for a n -photon Fock state is,

$$P_n(\alpha) = \frac{e^{|\alpha|^2}}{n!} \left(\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right)^n \delta(\alpha)\delta(\alpha^*) \quad (\text{A2})$$

We now combine (A1) with (A2) in (2) and integrate the expression to arrive at (3) of the main text,

$$\Pr(\bar{n}) = \frac{1}{\bar{n}!\sqrt{|\sigma_Q|}} \prod_j \left(\frac{\partial^2}{\partial\alpha_j\partial\alpha_j^*}\right)^{n_j} e^{\frac{1}{2}\alpha_v^\dagger A \alpha_v} \Big|_{\alpha=0}$$

We use Faà di Bruno's formula [36], a higher-order chain rule, to expand the derivatives in (3). Again, as in the main text, we restrict ourselves to $n_j = \{0, 1\}$ for N photons in total, meaning that we will have $2N$ derivatives remaining in (A3). The expansion of the derivatives [37] yields,

$$\frac{\partial^{2N} e^{\frac{1}{2}\alpha_v^\dagger A \alpha_v}}{\prod_i^N \partial\alpha_i \partial\alpha_i^*} = e^{\frac{1}{2}\alpha_v^\dagger A \alpha_v} \sum_{\pi} \prod_{B \in \pi} \frac{\partial^{|B|} \alpha_v^\dagger A \alpha_v}{\prod_{j \in B} \partial\alpha_j \alpha_j^{(*)}}, \quad (\text{A3})$$

where the sum is over all partitions π of the set of derivatives $\{\alpha_i^{(*)} = \alpha_i, \alpha_i^*\}$ and the product is the subgroups of this partition. One such partition, for two photons in modes 1 and 2, is $\{\alpha_1\}, \{\alpha_1^*, \alpha_2, \alpha_2^*\}$ which has the term,

$$\frac{\partial^{\frac{1}{2}} \alpha_v^\dagger A \alpha_v}{\partial\alpha_1} \frac{\partial^{\frac{3}{2}} \alpha_v^\dagger A \alpha_v}{\partial\alpha_1^* \partial\alpha_2 \partial\alpha_2^*}, \quad (\text{A4})$$

in the summation.

As $\alpha_v^\dagger A \alpha_v$ is a quadratic function in α_v all derivatives $\geq 3^{rd}$ order vanish and as the function is evaluated at $\alpha_v = 0$ all terms involving 1^{st} order derivatives also vanish. The final expansion is the $2N$ derivatives (whose indices are stored in μ) sorted into N sets of size 2, the PMP [17, 18]. The total number of PMP for $2N$ indices is $(2N-1)!!$, where $(.)!!$ is the double factorial. For example, the 3 PMP's of the numbers 1,2,3,4 are,

$$\{12\}\{34\}, \{13\}\{24\}, \{14\}\{23\}.$$

Applying this to exponential in (3) for two modes yields,

$$(A_{12} + A_{21})(A_{34} + A_{43}) + (A_{13} + A_{31})(A_{24} + A_{42}) \\ + (A_{14} + A_{41})(A_{23} + A_{32}).$$

where the indices 1,2 correspond to $\partial\alpha_1\partial\alpha_2$ and 3,4 to $\partial\alpha_1^*\partial\alpha_2^*$ i.e. the indices are i and $i+M$. As A is symmetric, $A_{i,j} = A_{j,i}$, and including the factor 1/2 leads to

$$\Pr(n_1=1, n_2=1) = \frac{A_{12}A_{34} + A_{13}A_{24} + A_{14}A_{23}}{\sqrt{|\sigma_Q|}} \\ = \frac{\text{Haf}(A)}{\sqrt{|\sigma_Q|}}. \quad (\text{A5})$$

The summation over all PMP's is the hafnian [19, 20]. The hafnian function is related to the better known pfaffian function [38] in the same way the permanent is related to the determinant, i.e by making all signs in the summation positive.

Appendix B: Multiple photons in the same mode

To deal with multiple photons in a single output mode, $n_j > 1$ in (3), we use the combinatorial method for partitioning sets with repetition [39]. This allows the problem to be written as the hafnian of an expanded matrix A'_S that is constructed by repeating entries of the original matrix.

An example: we have a single mode state, i.e. $M = 1$, described by the matrix A ,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (\text{B1})$$

and we detect two photons in the same mode, (mode 1)

$$\Pr(n_1 = 2) = \frac{1}{\sqrt{|\sigma_Q|}} \frac{1}{2!} \frac{\partial^2}{\partial\alpha_1^2} \frac{\partial^2}{\partial\alpha_2^{*2}} e^{\frac{1}{2}\alpha_v^\dagger A \alpha_v} \Big|_{\alpha=0}. \quad (\text{B2})$$

We now put one of the photons into a new mode, 2,

$$\Pr(n_1 = 1, n_2 = 1) = \frac{1}{2!\sqrt{|\sigma_Q|}} \frac{\partial}{\partial\alpha_1} \frac{\partial}{\partial\alpha_2} \frac{\partial}{\partial\alpha_1^*} \frac{\partial}{\partial\alpha_2^*} e^{\frac{1}{2}\alpha_v^\dagger A' \alpha_v} \Big|_{\alpha=0}, \quad (\text{B3})$$

where A' is the new matrix is given by,

$$A' = \begin{pmatrix} A'_{11} & A'_{12} & A'_{13} & A'_{14} \\ A'_{21} & A'_{22} & A'_{23} & A'_{24} \\ A'_{31} & A'_{32} & A'_{33} & A'_{34} \\ A'_{41} & A'_{42} & A'_{43} & A'_{44} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{11} & A_{12} \\ A_{21} & A_{22} & A_{21} & A_{22} \\ A_{11} & A_{12} & A_{11} & A_{12} \\ A_{21} & A_{22} & A_{21} & A_{22} \end{pmatrix}$$

Note that A' is **not** related to a proper quantum covariance matrix. It is just a tool to calculate the higher order derivatives. Also, the factor $\bar{n}!$ depends upon the original

distribution of photons and not the ‘new’ positions.

Appendix C: The Hafnian

The hafnian of a general, symmetric, matrix M can be decomposed into,

$$\text{Haf}(M) = \sum_{j=2}^{2N} m_{1,j} \text{Haf}(M_{[1,j]}) \quad (\text{C1})$$

where $M_{[1,j]}$ is the minor of the matrix M with the 1st and j^{th} rows and columns removed. Another decomposition useful for us is,

$$\begin{aligned} \text{Haf}(M) &= \text{Haf} \begin{pmatrix} A & B \\ B^t & C \end{pmatrix} \\ &= \text{Haf}(A) \text{Haf}(C) + \text{Perm}(B) \\ &+ \sum_{i_1[1,n], j_1[n+1,2n]}^{\binom{N}{1}^2} m_{i,j} \text{Haf}(M_{[i,j]}) \\ &+ \sum_{i_1, i_2[1,n], j_1, j_2[n+1,2n]}^{\binom{N}{2}^2} m_{i_1, j_1} m_{i_2, j_2} \text{Haf}(M_{[i_1, i_2, j_1, j_2]}) \\ &+ \dots + \\ &+ \sum_{i,j}^{\binom{N-1}{2}^2} m_{i_1, j_1} \dots m_{i_{N-1}, j_{N-1}} \text{Haf}(M_{[i_1, \dots, i_{N-1}, j_1, \dots, j_{N-1}]}) \end{aligned}$$

where other terms again are a summation of minors of the matrix M with rows/columns removed. As such we can relate the calculation of the hafnian to a summation of other hafnians and permanents, which are hard to approximate up to a multiplicative factor.

The physical significance of these terms is where the photons ‘originate’ from. The first term signifies all photons coming from the squeezed portion of the state, the second term all photons are thermal photons and the remaining terms represent different combinations i.e. 1 thermal photon, $N - 1$ squeezed photons (if $N - 1$ is even), 2 thermal photons, $N - 2$ squeezed photons etc.

Appendix D: Relationship between SBS and GBS

We will now demonstrate this relationship more rigorously by using the permanent and hafnian. We examine the situation of pumping every mode of the interferometer, $M \times \text{TMSS}$, with identical squeezing parameters r , generating $N (\leq M)$ photon pairs and for simplicity focus on a maximum one photon pair per TMSS.

For SBS,

$$\begin{aligned} \text{Pr}_{\text{SBS}}(\bar{n}|\bar{h}) &= \frac{|\text{Perm}(T_S)|^2}{\bar{n}! \bar{h}!} \\ &= \frac{\text{Perm}(T_S) \text{Perm}(T_S^*)}{\bar{n}! \bar{h}!}, \end{aligned} \quad (\text{D1})$$

where T_S is the submatrix of T , the matrix describing the interferometer, whose exact form is dependent upon the photon input and output.

Our Gaussian state at the output of the device has the covariance matrix,

$$\begin{aligned} \sigma &= \frac{1}{2} (\mathbb{I} \oplus T_{\text{SBS}} \oplus \mathbb{I} \oplus T_{\text{SBS}}^*) S \\ &\times S^\dagger (\mathbb{I} \oplus T_{\text{SBS}} \oplus \mathbb{I} \oplus T_{\text{SBS}}^*)^\dagger, \end{aligned} \quad (\text{D2})$$

where the order of the modes is $[\hat{a}_1 \dots \hat{a}_M, \hat{b}_1 \dots \hat{b}_M, \hat{a}_1^\dagger \dots \hat{a}_M^\dagger, \hat{b}_1^\dagger \dots \hat{b}_M^\dagger]$, with the \hat{a} -modes being the herald modes and the \hat{b} -modes are the T -transformed photons and S is the symplectic transformation that encodes the $M \times \text{TMSS}$,

$$S = \begin{pmatrix} \cosh r \mathbb{I}_{2M} & 0_M & \sinh r \mathbb{I}_M \\ 0_M & \sinh r \mathbb{I}_M & 0_M \\ \sinh r \mathbb{I}_M & 0_M & \cosh r \mathbb{I}_{2M} \end{pmatrix}. \quad (\text{D3})$$

The numerator of the right hand side of equ. (14) is,

$$\text{Pr}_{\text{GBS}}(\bar{n} \cap \bar{h}) = \frac{\text{Haf}(A_S)}{\bar{n}! \bar{h}! \sqrt{|\sigma_Q|}}. \quad (\text{D4})$$

It is straight-forward to show that the matrix A_S , with rows/columns deleted that have zero photons, has the form,

$$\begin{aligned} A_S &= -\tanh r \begin{pmatrix} 0 & T_S^\dagger & 0 & 0 \\ T_S^* & 0 & 0 & 0 \\ 0 & 0 & 0 & T_S^t \\ 0 & 0 & T_S & 0 \end{pmatrix} \\ &= B_S \oplus B_S^*. \end{aligned} \quad (\text{D5})$$

Using equ. (7),

$$\begin{aligned} \text{Haf}(A_S) &= \text{Haf}(B_S) \text{Haf}(B_S^*) \\ &= (\tanh r)^{2N} \text{Perm}(T_S) \text{Perm}(T_S^*) \\ &= (\tanh r)^{2N} |\text{Perm}(T_S)|^2. \end{aligned} \quad (\text{D6})$$

Finally, for $M \times \text{TMSS}$,

$$|\sigma_Q|^{-1/2} = \text{sech}^M(r). \quad (\text{D7})$$

Putting this all together,

$$\begin{aligned} \Pr_{\text{GBS}} &= \Pr(\bar{n} \cap \bar{h}) = \frac{\text{Haf}(A)}{\bar{n}! \bar{h}! \sqrt{|\sigma_Q|}} \\ &= \frac{\tanh^{2N}(r) |\text{Perm}(T_S)|^2 \text{sech}^{2M}(r)}{\bar{n}! \bar{h}!} \end{aligned} \quad (\text{D8})$$

and

$$\Pr(\bar{h}) = \text{sech}^{2M}(r) \tanh^{2N}(r), \quad (\text{D9})$$

thus

$$\frac{\Pr_{\text{GBS}}(\bar{n} \cap \bar{h})}{\Pr(\bar{h})} = \frac{|\text{Perm}(T_S)|^2}{\bar{n}! \bar{h}!} = \Pr_{\text{SBS}}, \quad (\text{D10})$$

as expected. As GBS can have any general passive transformation T , SBS is a subset of GBS.